



Wilson, A. T. (2018). Honesty as a Virtue. *Metaphilosophy*, 49(3), 262-280. <https://doi.org/10.1111/meta.12303>,  
<https://doi.org/10.1111/meta.12303>

Peer reviewed version

Link to published version (if available):

[10.1111/meta.12303](https://doi.org/10.1111/meta.12303)  
[10.1111/meta.12303](https://doi.org/10.1111/meta.12303)

[Link to publication record in Explore Bristol Research](#)  
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via WILEY at <https://onlinelibrary.wiley.com/doi/10.1111/meta.12303>. Please refer to any applicable terms of use of the publisher

## University of Bristol - Explore Bristol Research

### General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:  
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

# Honesty as a Virtue

Alan T. Wilson

**Abstract:** Honesty is widely accepted as a prime example of a moral virtue. And yet, honesty has been surprisingly neglected in the recent drive to account for specific virtuous traits. In paper, provides a framework for an increased focus on honesty by proposing success criteria that will need to be met by any plausible account of honesty. It then proposes a motivational account on which honesty centrally involves a deep motivation to avoid deception. It argues that this account satisfies the required success criteria, and discusses why (and when) honesty can be accepted as a moral virtue. Finally, it highlights some implications of accepting a motivational account of honesty for future work on encouraging the development of honesty, and on the relationship between different types of virtue.

Keywords: honesty, moral virtue, virtue ethics.

## Introduction

Honesty is widely accepted as an important moral virtue. And yet, honesty has been surprisingly neglected in the recent drive by virtue theorists to account for specific virtuous traits. This neglect is surprising, not only because of honesty's endorsement by influential virtue theorists, but also because of the trait's obvious relevance to current debates within (for example) politics, journalism, and sports. Efforts to encourage increased levels of honesty in these areas would be enhanced by having a clearer account of the nature of the trait.

In this paper, I aim to encourage an increased focus on the virtue of honesty. I will do this by, first, proposing four success criteria that will need to be met by any successful account (in Section II). Having surveyed Christian Miller's recent discussion of honesty (in Section III), I will then set out and defend my own motivational account of honesty, and demonstrate how

that account can satisfy the required criteria (in Sections IV and V). This will require a discussion of why (and when) honesty should be accepted as a moral virtue. Finally, I will highlight some implications that accepting a motivational account of honesty has for future work on the development of honesty, and on the relationship between different types of virtue (in Section VI).

## **I. The Importance of Honesty**

The claim that honesty is widely accepted as a moral virtue is likely to be met with little resistance. However, it is worth emphasising just how strong this acceptance has been. Influential virtue theorists have gone as far as to take for granted the virtue status of the trait. For example, in *Virtues of the Mind*, Linda Zagzebski (1996, 158) tells us that “honesty is on all accounts a moral virtue”, and that (Ibid., 86) “To think of virtue is almost immediately to think of examples of particular traits such as courage, generosity, compassion, justice, honesty, wisdom, temperance, and self-respect”. The same assumption in favour of honesty’s virtue status is made by Rosalind Hursthouse. For example, when testing her own naturalistic approach to identifying the virtues, Hursthouse (1999, Ch. 9) includes honesty as one of the traits that ought to be accommodated.

Importantly, the endorsement of honesty is not limited to theorists working in the Aristotelian tradition. For example, Julia Driver (2003), a consequentialist virtue theorist, dismisses rival attempts to identify the moral virtues partly because of their supposed failure to accommodate honesty. And Michael Slote (2011) uses an assumption in favour of the virtue status of honesty when arguing for his controversial claim that some virtues necessarily conflict with one another. Given this widespread endorsement of honesty, there is good reason for virtue theorists to explore the nature of the trait.

Further reason for exploring the nature of honesty is provided by current debates within politics, journalism, and sports (to mention just three examples). The demand for honesty from our politicians (and the suspicion that honesty is not always forthcoming) may be nothing new. But the use of social media appears to have exacerbated concerns about whether politicians and the media can be trusted to engage truthfully with the electorate. Such concerns have been well-documented in recent times, as have concerns about levels of dishonesty and manipulation in the sporting world. The credibility of many sporting contests has been damaged by a series of recent scandals involving dishonesty and cheating, be that through illicit use of performance-enhancing substances, or through more mundane methods, such as match-fixing (tennis); diving and play-acting (football); using fake blood capsules to feign injury (rugby); or even by insufficiently inflating footballs (American football). In all of these areas, there is a suspicion that participants are not being honest, and a concern that the very legitimacy of these enterprises may be damaged as a result of (the perception of) rampant dishonesty.

And yet, the candidate virtue of honesty has been almost entirely neglected by virtue theorists. This is true despite a recent move in virtue theory away from attempting to give an over-arching account of virtue *in general*, and towards providing accounts of *specific* virtuous traits.<sup>1</sup> If we believe that honesty is important, and if we want to correct for a lack of honesty (either in ourselves or more generally), then it will be vital to have an account of the nature of the trait. In the next section, I begin the task of encouraging an increased focus on honesty by generating success criteria that could be used to assess any proposed account of the trait.

---

<sup>1</sup> See, for example, Timpe and Boyd (2014), Battaly (2010, Part 4), and Roberts and Wood (2007, Part II). The neglect of honesty is illustrated by the fact that Timpe and Boyd's important collection on virtues and vices not only fails to have a chapter dedicated to honesty, but honesty does not even appear once in the book's index.

## II. Generating Success Criteria

Success criteria for an account of honesty can be generated by considering why some initial attempts to explain the trait are unsuccessful. For example, a basic starting point when attempting to get clearer about honesty would be to say that an honest agent is one who does not tell lies. But this will not do, because we can imagine an agent who does not tell lies, but who does not deserve to be considered honest either. For example, a Robinson Crusoe figure who never interacts with another human being, and so never tells a lie, would not thereby deserve to be considered honest.<sup>2</sup> This worry cannot be avoided simply by amending the initial basic claim and saying that an honest agent is one who does not tell lies *and* who regularly utters truths. We can imagine an agent who simply repeats everything that she overhears, and who never happens to overhear a falsehood. Such an agent is not necessarily honest.

A slightly modified starting point would be to say that an honest agent is one who does *whatever is appropriate* with regards to either lying or telling the truth. One worry about the example of someone who merely repeats whatever she overhears might be that such a person is not properly responsive to whether it would be appropriate to tell the truth. But this modified approach is also problematic for (at least) two reasons. Firstly, the approach faces a worry concerning occasions where the appropriate act is to tell a lie (or to conceal a truth). While we might want to say that lying on these occasions is *compatible* with being virtuously honest, it would nevertheless be strange to say that the lie thereby *expresses* the

---

<sup>2</sup> A Robinson Crusoe example is used by Driver (2001, Ch. 2) to reject a similarly basic account of modesty.

agent's honesty. And yet, if honesty just is a disposition to do whatever is appropriate with regards to the truth, then that is the conclusion that we would have to accept.<sup>3</sup>

A second, and more damaging, worry for any approach that focuses solely on lying (or on telling the truth) is that such accounts are too narrow. They miss out on the full range of what is involved in being honest. As has been noted elsewhere, there is more to honesty than an appropriate disposition towards telling the truth. For example, when discussing what we would expect from an honest agent, Hursthouse (1999, 10) provides the following list:

they do not lie or cheat or plagiarize or casually pocket other people's possessions.

You can rely on them to tell you the truth, to give sincere references, to own up to

their mistakes, not to pretend to be more knowledgeable than they are; you can

buy a used car from them or ask for their opinion with confidence.

Hursthouse is not alone in recognising the different ways in which honesty is manifested.<sup>4</sup> In a recent contribution, Christian Miller helpfully groups together behaviours associated with honesty into five seemingly distinct categories. The categories identified by Miller (2017, 240) can be paraphrased as follows:

*Truthfulness*: Being disposed to reliably tell the truth.

*Being Respectful of Property*: Being disposed to reliably respect the property of others.

---

<sup>3</sup> This worry is raised by Gary Watson for what he calls the *due concern* view of virtue. Having set out the implication, Watson (1984, 68) points out that "no one would think to say such a thing".

<sup>4</sup> See also Kamtekar (2004, 468-469).

*Proper Compliance*: Being disposed to reliably follow the relevant rules.

*Fidelity to Promises*: Being disposed to reliably keep reasonable promises.

*Forthrightness*: Being disposed to reliably give a complete presentation of the facts.

Miller (Ibid., 240-241) refers to these as “virtues” or “subvirtues” of honesty, and he is open to the possibility that additional subvirtues could be uncovered. To avoid complications, I will avoid the terminology of subvirtues and instead refer to these as five possible *aspects* of honesty. I will not consider whether each aspect ought to also be considered virtuous in its own right. The important point is that honesty is plausibly connected to (at least) these different forms of behaviour, and so an account of honesty that focuses solely on telling the truth and avoiding falsehoods is likely to be inadequate.

These considerations present us with our first success criterion for an account of honesty. Immediately after identifying these five aspects, Miller (Ibid.) proposes what he calls the “unification challenge”. This is the challenge of explaining why honesty is thought to be manifest in seemingly distinct ways. I suggest that an initial criterion for any successful account of honesty is that it be able to meet Miller’s unification challenge. That is, a successful account of honesty will shed light on why honesty is closely associated with a range of seemingly distinct forms of behaviour, including the five aspects of honesty listed above.

Basic accounts on which honesty is simply a matter of avoiding lies or telling the truth fail to satisfy this initial criterion. They cannot explain why an honest agent is also expected to behave in other ways, such as keeping promises, and avoiding any temptation to pocket the possessions of others. But these basic accounts have also been shown to fail in another way. They result in implausible verdicts when applied to cases such as the Robinson Crusoe

example, or the example of someone who just happens to utter exclusively true statements. This reveals a second success criterion. In addition to explaining why the five aspects above are relevant to the single trait of honesty, a successful account will also lead to plausible verdicts concerning who is or is not an honest agent.

The importance of these initial success criteria can be shown by surveying two of the few discussions of honesty as a virtue that can be found in the literature. David Carr (2014, 9) argues that “the soul of honesty as a moral virtue” is a “disposition to seek and honor truth”. Does this account meet Miller’s unification challenge, at the same time as providing plausible verdicts about who is or is not honest? It is possible that Carr’s account meets the unification challenge. This will depend on what it means to appropriately “honor” the truth. Rather than pursue that question here, we can instead see that Carr’s account fails to satisfy the second criterion of generating plausible verdicts. The claim that honesty centrally involves a disposition to “seek and honor” truth is too broad. An agent who is not motivated to seek out new truths might nevertheless avoid being dishonest, as long as she is generally forthright with the information that she does possess. Carr (Ibid.) mentions that, on his account, honesty bears some similarity to curiosity and open-mindedness. I think this is right, but it gives us good reason to question whether we have uncovered the correct account of *honesty*. By implying that an agent who lacks inquisitiveness must thereby be dishonest (or at least lacking in honesty), this proposal fails to satisfy the second of the two criteria that have been proposed so far.

Another rare extended discussion of honesty as a virtue is provided by Louis Geunin, although Geunin’s focus is narrower, in that he is primarily concerned with *intellectual* honesty. According to Geunin (2005, 180), intellectual honesty is but one component of honesty in general, and is concerned with “truthfulness and veraciousness”. These are



understood, respectively, as “the attribute of being truthful” and “the disposition to be truthful”. Honesty more generally is explained (Ibid., 222) as “an inclusive disposition embracing veraciousness as well as dispositions not to cheat and not to violate rules of fair play. As it is said, one who is honest does not lie, cheat, or steal.” It is apparent that this account does not meet the unification challenge. When discussing honesty in general, Geunin recounts the different behavioural dispositions that we would expect from an honest agent, but does not offer an explanation of why those different aspects tend to hang together. This should not be taken as a criticism of Geunin, whose aim is not to provide any such explanation. But it does demonstrate that we will need to look elsewhere for an account of honesty in general that satisfies the initial criteria.

Before moving on, I want to suggest two further success criteria. While the virtue of honesty has been relatively neglected in the virtue literature, this is not the case for all candidate moral virtues. For example, the trait of modesty has been subject to a high level of scrutiny and debate.<sup>5</sup> While I will not discuss that debate here, the literature on modesty can provide inspiration for additional success criteria when assessing accounts of specific virtuous traits.

The flourishing debate concerning modesty is due, in large part, to an account proposed by Julia Driver. When setting out her account, Driver (2001, Ch. 2) argues that it is capable of doing two things that any account of modesty ought to achieve. First, she claims that her account is compatible with a plausible corresponding account of the vice of false modesty. Second, she argues that her account can explain the value or virtuousness of (genuine) modesty. Regardless of whether Driver’s account is, in fact, successful in these ways, these

---

<sup>5</sup> See, for example, Driver (2001, Ch. 2); Ridge (2000); McMullin (2010); and Wilson (2016).

do appear to be relevant aims when proposing an account of a specific virtuous trait. I therefore suggest the following additional success criteria for an account of honesty: First, the account ought to be compatible with a plausible corresponding account of dishonesty. Second, the account ought to go some way towards explaining why honesty is widely accepted as a moral virtue. Adding these two criteria to those considered above results in the following list of what we should expect from any plausible account of honesty:

*The Unification Challenge:* An account of honesty ought to explain why the trait is thought to be manifest in a range of seemingly distinct behaviours, including the five aspects of honesty noted earlier.

*Plausible Verdicts:* An account of honesty ought to produce plausible verdicts concerning who should (or should not) be classed as an honest agent.

*Corresponding Account of Dishonesty:* An account of honesty ought to be compatible with a plausible corresponding account of dishonesty.

*Explanation of Value:* An account of honesty ought to at least point in the direction of an explanation for why honesty is a moral virtue.

By setting out these four success criteria, I aim to provide a framework for further discussion of honesty as a virtue. The challenge, then, is to offer an account of the nature of honesty that can satisfy these criteria. In the remainder of this paper, I will take up this challenge by proposing and defending a motivational account of honesty, and by explaining what this account has to say about the virtue status of the trait.

### III. Miller's Starting Point

In order to provide an account of honesty that can satisfy all four of the proposed success criteria, it is necessary to first meet Miller's unification challenge. This requires explaining what lies at the core of honesty, such that honesty would be expected to be manifest in the seemingly distinct ways identified by the five aspects of honesty that have been identified – *truthfulness*; *being respectful of property*; *proper compliance*; *fidelity to promises*; and *forthrightness*. Given the aim of meeting this challenge, it is worth considering Miller's own proposal.

The initial proposal from Miller (2017, 244) is as follows:

The virtue of honesty is, centrally, a character trait concerned with reliably not intentionally distorting the facts.

Miller (Ibid., 244) is clear that he views this mainly as "a promising starting point" aimed at encouraging future work. We can see both why the account is promising and why it will require revision by assessing it in terms of the success criteria generated in the previous section.

Does this proposal meet the unification challenge? Miller (2017, 241-244) first explains how the account can accommodate four of the five aspects of honesty mentioned above, before then discussing more problematic cases. It is plausible that *truthfulness* will involve not intentionally distorting the facts, as will *proper compliance* (such as by avoiding match-fixing or feigning injury in sports) and *being respectful of property* (such as by avoiding passing off another's possessions as one's own). We can also see how not intentionally distorting the facts is connected to *forthrightness*. An agent who is less than forthright might

typically be hoping to induce a false understanding of the facts, either through implication or by omission. In these ways, Miller's initial proposal can explain why honesty is associated with four of the five aspects of honesty that were highlighted above.

The first of the more problematic cases discussed by Miller (*Ibid.*, 242) is *fidelity to promises*. In what way does the breaking of a promise involve an intentional distorting of the facts? First of all, we can see why making false promises is associated with (dis)honesty. By promising to repay the money despite having no intention of doing so, I distort facts about what I can be expected to do in future (and about my current intentions). But what about cases where an agent genuinely intends to keep a promise, but later fails to do so? Is breaking a promise in this way a sign of dishonesty, and, if so, can it be understood as an intentional distorting of the facts? The correct response to this issue is to recognise that not all failures to keep a promise indicate a lack of honesty.

*Fidelity to promises* is properly associated with honesty, but only because a consistent failure to keep one's promises will justifiably raise suspicions about whether those promises are being made in good faith. If they are not, then one is being dishonest, and Miller's initial proposal is able to accommodate this. But the breaking of a promise does not always indicate a lack of honesty, such as when the promise is made in good faith, but cannot ultimately be fulfilled due to unanticipated complications. In these cases, an agent might reveal a failure of character, but this is likely to be a failure of due care, or perhaps a failure of justice. Such cases do not necessarily reveal a lack of honesty.<sup>6</sup> Given this, we can explain why *fidelity to promises* is rightly associated with honesty, even if not all cases of breaking a promise should be thought of as dishonest.

---

<sup>6</sup> Miller (*Ibid.*) suggests that such cases instead indicate a lack of loyalty.

A second problematic case discussed by Miller (Ibid., 243) can be responded to in the same way. Even if the initial proposal does, in general, explain why honesty is associated with *being respectful of property*, there might be cases where stealing another's property involves no obvious distortion of the facts. Miller's example is of a "bald-faced theft" where the perpetrator is open about what he is doing, and does not pretend that the stolen property belongs to him. Even if this is an example of lacking respect for property that does not involve any obvious distortion of facts, Miller is correct to suggest that this case is not necessarily one that involves dishonesty. As with the previous problematic example, we can say that an agent who engages in bald-faced theft reveals a failing of character. But this failure will be one of justice, or of charity, and not necessarily a failure of honesty. I believe that this is an intuitively plausible verdict in response to cases of bald-faced theft, and so Miller's initial proposal is also able to accommodate those instances of *being respectful of property* that are properly associated with honesty.

And so, having set up the unification challenge, Miller then provides an initial proposal that is capable of meeting that challenge. This demonstrates the promise of the proposal. That the proposal should only be thought of as a starting point is demonstrated by how it fares in relation to the second criterion generated earlier: the need to produce plausible verdicts concerning who should (or should not) be classed as an honest agent.

One worry for the proposal, as it stands, is that it provides only a negative requirement for being an honest agent. By this I mean that the proposal highlights a reliable *lack* of certain behaviours (intentional distortions of the facts) as the central feature of honesty. And yet, we can imagine an agent who has this central feature but who is not necessarily honest. For example, an agent suffering from depression, or from a general apathy, might cease to care about the facts one way or another, and so be reliably disposed not to bother intentionally

distorting any facts (at the same time as being reliably disposed not to bother checking any facts either). Such an agent appears to satisfy Miller's initial proposal, but it is not clear that this is enough to count as being a possessor of genuine honesty.

In addition to worries about the central feature of honesty being presented negatively (as a lack), there is a further worry that Miller (*Ibid.*, 245) identifies. It is possible for an agent to be reliably disposed not to distort the facts, but where this reliability has a positive explanation that is somehow inappropriate. For example, someone might reliably not distort the facts only because of an active desire to avoid punishment.<sup>7</sup> If we think that this underlying explanation casts doubt on an agent's possession of honesty (or of morally virtuous honesty), then this will be problematic for Miller's initial proposal.<sup>8</sup>

Miller is aware that his initial proposal may need to be amended. For this reason, he mentions (*Ibid.*, 244-246) four possible revisions to his proposal, while being clear that his goal "is not to argue either for or against these revisions" in his paper. The possible revisions mentioned by Miller are as follows:

- (1) The virtue of honesty is, centrally, a character trait concerned with reliably not distorting the facts *unless it is morally appropriate to do so*.
- (2) The virtue of honesty is, centrally, a character trait concerned, *for good or virtuous motivating reasons*, with reliably not intentionally distorting the facts.
- (3) The virtue of honesty is, centrally, a character trait concerned with reliably not intentionally distorting the facts *as the agent sees them*.

---

<sup>7</sup> Miller's own examples are of seeking to make a good impression, and seeking rewards in the afterlife.

<sup>8</sup> Similar worries may arise for Smith's (2003, 518) account on which honesty is "a person's refusal to pretend that facts are other than they are".

- (4) The virtue of honesty is, centrally, a character trait concerned with reliably not intentionally distorting the facts *in one's own life and also actively preventing such distortions in others (in morally appropriate ways)*.<sup>9</sup>

It will not be possible to evaluate the four potential revisions to his initial proposal that Miller mentions. However, given the problem cases just discussed, I believe that the second potential revision is the most promising. The reliability of an agent suffering from depression, or of an agent seeking to avoid punishment, is plausibly not explained by any “good or virtuous motivating reasons”, and so those agents will no longer satisfy Miller’s account once it is revised in this way. One way of attempting to advance the debate on honesty, therefore, would be to develop this second potential revision by providing arguments in favour of it (perhaps by arguing that it can meet the four success criteria set out above) and by providing an account of how the “good and virtuous” motivating reasons are to be understood.

And yet, that will not quite be my approach here. I believe that Miller’s discussion provides a helpful starting point for the debate, particularly when demonstrating how an account might meet his unification challenge. I am also generally sympathetic towards accounts of virtue that emphasize an important role for an agent’s motivations, as is suggested by Miller’s second potential revision. However, as we shall see, I believe that an agent’s motivation needs to play two distinct roles (or a role at two distinct levels) when providing an account of the nature of honesty as a moral virtue.

Firstly, I believe that an honesty-specific motivation ought to be identified as the central feature of honesty in general (rather than a tendency to reliably act, or to avoid acting,

---

<sup>9</sup> Emphasis added.

in certain ways). Only then should a further appeal be made to the idea of virtuous motivations in order to explain when and why honesty is rightly considered morally virtuous. In what follows, I want to propose and defend an account that has these features. This will involve proposing both what the honesty-specific motivation could be, and how we might think of the additional virtuous motivations that are present when honesty is a moral virtue. I will argue in favour of this account by demonstrating that it can satisfy the success criteria outlined above.

#### **IV. A Motivational Account of Honesty**

The account of honesty in general that I want to propose can be set out as follows:

*The Motivational Account of Honesty:* The trait of honesty centrally involves a deep motivation to avoid deception.

This account can be categorized as a *motivational* account of honesty for the simple reason that it identifies a motivation as the central feature of honesty, as opposed to a reliable behavioural disposition, or a tendency to produce certain outcomes. On this account, honesty requires a deep motivation to avoid deception. What is meant by a “deep” motivation in this context? We should not accept that just any motivation to avoid deception is sufficient for honesty. Instead, an agent’s motivation must have certain features.<sup>10</sup> Firstly, the motivation must be sufficiently *persistent*, in the sense that it is neither fleeting nor sporadic. Secondly, the motivation must be sufficiently *strong*, in the sense that it is capable of actually

---

<sup>10</sup> See also Wilson (forthcoming).



influencing the agent's behaviour. And thirdly, the motivation must be sufficiently *robust*, in the sense that it is not easily defeated by competing considerations. We would not consider an agent to be honest if her persistent and strong motivation to avoid deception was consistently overridden by her competing desire to be well-liked. Only when an agent's motivation has these three features can we say that she possesses a *deep* motivation to avoid deception, and only then can she be considered truly honest.

This account of honesty can respond to the unification challenge. An agent who is deeply motivated to avoid deception can be expected to avoid telling lies and to avoid making false promises, and this explains why honesty is associated with *truthfulness* and *fidelity to promises*. Similarly, an agent who is deeply motivated to avoid deception will avoid presenting only a partial and misleading picture of the truth, and so honesty can also be expected to result in *forthrightness*. In this way, the motivational account explains the connection between the trait of honesty and three of the five aspects of honesty mentioned above.

The remaining two aspects of honesty can be accommodated in a similar way. Efforts to cheat or to violate the rules of fair play typically involve deception. This is true, for example, when a footballer pretends to have been unfairly challenged by an opponent, or when a tennis player pretends to be trying to win her match despite having secretly agreed to lose in exchange for a cash payment. An agent who is deeply motivated to avoid deception would not be expected to act in these ways, and this explains why honesty is associated with *proper compliance*. And it is also true that most cases of theft also involve deception. This is the case, for example, when an agent poses as a security guard in order to rob a bank, or when claiming stolen possessions as one's own after the fact. Again, an agent who was motivated to avoid deception would not be expected to act in these ways, and this explains why honesty is associated with *being respectful of property*. By accommodating all five of the aspects of

honesty that have been identified, the motivational account provides a satisfying response to Miller's unification challenge.

There may be cases where an agent fails to demonstrate either *fidelity to promises* or *being respectful of property* without engaging in any deception. These cases will include the problematic examples discussed by Miller, and considered above. An agent who makes a promise in good faith might later fail to fulfil that promise due to unforeseen complications while not engaging in deception of any kind. An agent who is guilty of bald-faced theft when stealing from others might fail to be respectful of property while not engaging in deception of any kind. If this is possible, then these actions will be compatible with a deep motivation to avoid deception, and so will be judged as compatible with honesty on the proposed account. And yet, this is not a problem for the account. We saw above that the failings involved in these problematic cases are better viewed as failures of justice, or of charity, rather than as failures of honesty. That verdict is compatible with the motivational account of honesty. Therefore, we can maintain our initial verdict in response to these problematic cases, and say that the motivational account is able to explain why the five aspects of honesty are all properly associated with the single trait of honesty.

The motivational account is also able to produce plausible verdicts in response to examples that proved problematic for rival accounts. Neither Robinson Crusoe nor someone who merely repeats whatever she overhears will be considered honest on this account. This is because their behaviour (of avoiding falsehoods and uttering truths) is not explained by a deep motivation to avoid deception. And unlike Carr's account, the motivational account does not demand the seeking out of new information in order for someone to be considered honest. If an agent is not motivated to uncover exciting new truths, then this might tell against her possession of an important virtue (such as inquisitiveness or intellectual curiosity). But it

will not tell against her possession of honesty, as long as she does have a deep motivation to avoid deception. In these ways, the motivational account is better able to meet the second criterion of generating plausible verdicts than were proposals considered above.

The motivational account can also satisfy the third criterion of being compatible with a plausible corresponding account of dishonesty. According to the motivational account, dishonesty consists in a problematic failure to be deeply motivated to avoid deception.<sup>11</sup> The most extreme way of failing in this regard is to have a positive motivation to deceive. Any agent motivated in this way clearly deserves to be considered dishonest. And yet, we might expect that such agents will be rare. More common will be those who lack honesty as a result of having a motivation to avoid deception that falls below the necessary levels of persistence, strength, or robustness. Someone who deceives others because her motivation to avoid deception is always overridden by her competing desire to be well-liked ought to be considered dishonest. Someone who deceives others because her motivation to avoid deception is problematically weak or sporadic ought to be considered dishonest. The understanding of dishonesty that follows from the motivational account of honesty gains plausibility by securing these results.

In these ways, the proposed motivational account of honesty satisfies three of the four success criteria. It can explain why honesty is thought to be manifest in seemingly distinct behaviours; it generates plausible verdicts concerning who is and who is not honest; and it is compatible with a corresponding account of dishonesty that is intuitively appealing. Can the proposed account also explain the moral virtuousness of honesty, thereby satisfying the final

---

<sup>11</sup> Some will fail to possess a sufficiently deep motivation in ways that are not problematic. Very young children, for example, might lack a sufficiently deep motivation to avoid deception and yet not be properly thought of as dishonest.

criterion? I do not believe that the account is sufficient, as it stands, to achieve this result. In the next section, I will explain why this is the case, before then detailing an amendment to the account that allows it to satisfy all four of the proposed success criteria.

## V. Honesty as a Moral Virtue

Approaches to identifying which traits are morally virtuous can be separated into two broad categories. On an *outcomes-based* approach, the moral virtues are those aspects of an agent that reliably lead to positive outcomes. On a *motivations-based* approach, the moral virtues are those aspects of an agent that involve intrinsically valuable motivations.<sup>12</sup> Different theorists then disagree about which are the relevant positive outcomes, and about which motivations are intrinsically valuable. The challenge for any account of honesty, then, is to demonstrate that honesty (on the proposed understanding) will be accepted as a moral virtue by one, or both, of these two approaches. I will here focus on whether honesty would be accepted as a virtue by a motivations-based approach to identifying moral virtues.<sup>13</sup>

I have argued that honesty centrally involves a deep motivation to avoid deception. If this motivation is accepted as having intrinsic moral value, then honesty will be accepted as a moral virtue by a motivations-based approach to identifying moral virtues. In that case, the proposed account will already satisfy all four of the required success criteria. However, it is not clear to me that the characteristic motivation of honesty does deserve to be included as

---

<sup>12</sup> For more on these two approaches, see Battaly (2015).

<sup>13</sup> This may actually be the more difficult task, given the general acceptance that honesty is beneficial in producing positive outcomes such as well-functioning social groups, and intimate relationships. See, for example, Hursthouse (1999, 209-210) and LaFollette and Graham (1986).

being of fundamental or intrinsic moral value. It seems possible to possess a deep motivation of this sort in a way that is not morally virtuous.

When discussing Miller's proposal, I used the example of someone who reliably does not intentionally distort the facts, but only because of a fear of being punished. Such an agent is not obviously morally virtuous. However, we can also imagine an agent who is deeply motivated to avoid deception only because of a further, more fundamental, motivation to avoid punishment. Even if such an agent should be accepted as being honest, this is not an obvious case of moral virtue. For this reason, the proposed motivational account of honesty requires additional support so as to properly explain when honesty should be accepted as morally virtuous. The necessary amendment can be explained by first taking inspiration from an influential account of the intellectual virtues.

In *Virtues of the Mind*, Linda Zagzebski (1996, Part II) sets out an understanding of the intellectual virtues that appeals to a distinction between proximal motivations and underlying motivations. A proximal motivation is one that is grounded in a more fundamental, underlying motivation. For example, I might have a proximal motivation to save a percentage of my income every month, where this proximal motivation stems from an underlying motivation to live well in retirement. Crucially, Zagzebski argues that intellectual virtues involve both a proximal motivation and an underlying motivation. First of all, each intellectual virtue has its own characteristic proximal motivation. For example, the virtue of intellectual rigour might involve the characteristic proximal motivation to thoroughly examine evidence, while the virtue of intellectual courage involves the characteristic proximal motivation to respond fearlessly to intellectual challenges.<sup>14</sup> But in the case of intellectual virtues, these proximal

---

<sup>14</sup> For more examples, see Zagzebski (1996, 269).

motivations are grounded in a shared underlying motivation – the motivation to achieve “cognitive contact with reality”.<sup>15</sup> And it is this shared underlying motivation that explains why they are intellectually virtuous. I believe that Zagzebski’s idea of proximal and underlying motivations can be used to amend the proposed account of honesty so as to explain the status of honesty as a moral virtue.

In addition to an account of the nature of honesty, it is also important to explain why (and when) honesty is morally virtuous. The account that I have in mind is as follows:

*The Motivational Account of Honesty as a Moral Virtue:* Honesty is morally virtuous only when the characteristic motivation to avoid deception is grounded in an underlying motivation that is of intrinsic moral value.

Different theorists can accept this account, even if they then disagree about which motivations have intrinsic moral value. It will not be possible here for me to argue in favour of any particular position on this issue. However, I have argued elsewhere that fundamentally valuable moral motivations include the motivation to protect and promote well-being, and the motivation to ensure fairness.<sup>16</sup> Just as intellectual rigour is an intellectual virtue when its characteristic motivation to examine the evidence is grounded in an underlying motivation for “cognitive contact with reality”, so too will honesty be a moral virtue when its characteristic motivation to avoid deception is grounded in, for example, an underlying motivation to ensure fairness.

---

<sup>15</sup> For explanation of this term, see *Ibid.*, 131-132. I also provide a slightly more detailed summary of Zagzebski’s account in Wilson (forthcoming).

<sup>16</sup> See Wilson (forthcoming).

Using the motivation to ensure fairness as an example of a morally valuable motivation, we can see how the proposed account can lead to plausible verdicts concerning when honesty is morally virtuous. We should not say that an agent who is deeply motivated to avoid deception only because of a fear of being punished is morally virtuous. Even if this agent is properly considered honest (if the motivation is sufficiently deep), she should not be considered *virtuously* honest. A deep motivation to avoid deception is only morally virtuous when it is grounded in an intrinsically valuable underlying motivation, such as the motivation to ensure fairness. An agent who is deeply motivated to avoid deception because she thinks that it would be unfair to deceive others can be both honest and morally virtuous. An agent who is deeply motivated to avoid deception only because she is afraid that she will be punished may well be honest, but she will not thereby be morally virtuous.

Combining the initial motivational account of honesty with this further account of when honesty is a moral virtue renders the overall account better able to explain why (and when) honesty can be included as a moral virtue. As a result, the motivational account becomes able to satisfy all four of the success criteria that were generated in Section II. By explaining how these criteria can be satisfied, I hope to have demonstrated the promise and plausibility of the proposed account.

## **VI. Further Implications**

Before concluding, I want to briefly discuss two implications of my proposal. The aim of this section is to highlight possible areas of future research for those interested in honesty as a virtue.

## VI.I. ENCOURAGING THE DEVELOPMENT OF VIRTUOUS HONESTY

The claim that honesty centrally involves a characteristic motivation, as opposed to a characteristic behavioural disposition, has important implications concerning how to encourage the development of honesty. If our aim is only to encourage certain behaviours, then it might be thought effective to provide rewards for the desired behaviour, and to enforce punishments for any undesired behaviour. This is a strategy that could be suggested in an educational context. For example, a school might include a selection of “virtue awards” at the annual prizegiving ceremony in order to reward virtuous behaviour, or even choose to amend grades on the basis of students’ perceived virtues (or vices). In these ways, parents and educators might choose to use rewards and/or punishments when viewing the aim of virtue education as being to encourage virtuous behaviour.

However, this will not be an effective strategy for encouraging the development of honesty as a virtue, once we accept that honesty centrally involves a deep motivation to avoid deception. This is because there is evidence that offering external rewards (and threatening external punishments) may actually have a negative impact on motivational development. As Heather Battaly (2015, 155) explains:

There is mounting evidence that extrinsic motivation for tangible rewards (e.g., money, awards) does not facilitate, and even undermines, intrinsic motivation. According to Edward Deci, Richard Koestner, and Richard Ryan, a host of psychological studies have shown that offering students tangible rewards for tasks



like reading and writing actually *decreases* their intrinsic motivation to perform such tasks.<sup>17</sup>

If this is correct, and if honesty centrally involves a deep motivation to avoid deception, then it will be important to *avoid* offering rewards or punishments in response to behaviour that is associated with honesty (such as the behaviours identified by the five aspects of honesty). Such incentives to act honestly have the potential to decrease an agent's characteristically honest motivation, and so make it more difficult for virtuous honesty to develop.

The evidence discussed by Battaly suggests that existing strategies used by parents and educators when attempting to inculcate honesty may be self-defeating if our aim is to encourage virtuous honesty. This has implications for work that will need to be carried out by those interested in encouraging the development of honesty. Firstly, it will be important for virtue theorists to engage with psychological studies, with the aim of either confirming or refuting the negative impact of rewards and punishments on motivational development. Secondly, it will also be important for more work to be done on uncovering and assessing alternative methods for encouraging the development of virtuous honesty. This will include focusing on strategies that do not involve obvious external incentives. It is possible that such strategies will include role-modelling; the use of art and literature; and perhaps educating children directly about virtue concepts.<sup>18</sup>

---

<sup>17</sup> The study mentioned by Battaly is Deci, Koestner, and Ryan (2001).

<sup>18</sup> A list of seven methods for encouraging intellectual virtues in the context of education is provided by Jason Baehr (2013, 256-259).

## VI.II. OTHER FORMS OF VIRTUOUS HONESTY

A second implication of the motivational account concerns the possibility of forms of honesty that are virtuous, but that are not morally virtuous. A key feature of the motivational account is that it separates the explanation of the central feature of honesty (the deep motivation to avoid deception) and the explanation of why and when honesty is morally virtuous (the grounding of the honest motivation in a morally valuable underlying motivation). This has implications for the virtue status of honesty.

As discussed above, the motivational account makes it possible for someone to be properly considered honest (by having the required deep motivation) while at the same time not possessing honesty as a virtue (if the required motivation is grounded in an underlying motivation that is somehow inappropriate). However, the account also leaves open the possibility that someone could possess honesty as a non-moral virtue. For example, the idea that honesty might be better understood as an intellectual virtue has been discussed elsewhere.<sup>19</sup> The motivational account makes it possible to understand how this could occur.

Honesty is a moral virtue when the motivation to avoid deception is grounded in a morally valuable underlying motivation, such as the motivation to protect and promote well-being, or the motivation to ensure fair outcomes. But it may also be possible for the motivation to avoid deception to be grounded in an underlying motivation that is intellectually valuable, or perhaps aesthetically valuable. Honesty might be possessed as an intellectual virtue, such as when the motivation to avoid deception is grounded in a love of truth. Or honesty might be possessed as an aesthetic virtue, such as when the motivation to avoid deception is grounded in a desire for artistic authenticity, or in a belief that the truth is

---

<sup>19</sup> See Driver (2003, 381) and Carr (2014, 4). Intellectual honesty is the focus of Guenin (2005).

somehow beautiful. The opening up of these possibilities ought to encourage virtue theorists to explore the idea that familiar candidate virtues are not necessarily restricted to one particular category of virtue. Instead, future work on honesty might allow for a more open-minded discussion about the exact relationship between a range of different types of virtue, including moral, intellectual, and aesthetic virtues.<sup>20</sup>

In these ways, the motivational account of honesty points towards important future work at both a practical and a conceptual level for those interested in honesty as a virtue. It is hoped that this paper will encourage such work, and that it will result in an increased engagement by virtue theorists with the neglected virtue of honesty.

## **Conclusion**

In this paper, I have attempted to address the current neglect of honesty by virtue theorists. I have provided a framework for future discussions by generating four success criteria that can be used to assess proposed accounts of honesty. And I have proposed a motivational account of honesty, arguing that this account satisfies the required criteria. It is hoped that this will encourage others to join in the task of accounting for the nature of honesty, and that it will prompt future research on how best to encourage the widespread development of honesty as a virtue.

*Department of Philosophy*

*Cotham House*

*University of Bristol*

---

<sup>20</sup> I take steps in this direction in Wilson (forthcoming).

Bristol

United Kingdom

BS6 6JL

alan.wilson@bristol.ac.uk

## Acknowledgements

For helpful comments and discussion, I am grateful to Christian Miller, participants at the Aretai Conference in Genoa, and to the anonymous reviewers. I am also grateful to Michel Croce and Maria Silvia Vaccarezza for inclusion in this special issue. Work on this paper was supported by a grant from the Templeton Religion Trust. The opinions expressed here are those of the author, and do not necessarily reflect the views of the Templeton Religion Trust.

## References

- Baehr, Jason. 2013. "Educating for Intellectual Virtues: From Theory to Practice." *Journal of Philosophy of Education*, vol. 47, no. 2, pp. 248-262.
- Battaly, Heather. 2010. *Virtue and Vice, Moral and Epistemic*. Chichester: Wiley-Blackwell.
- 2015. *Virtue*. Malden, MA: Polity Press.
- Carr, David. 2014. "The Human and Epistemic Significance of Honesty as an Epistemic and Moral Virtue." *Educational Theory*, vol. 64, no. 1, pp. 1-14.
- Deci, Edward L., Richard Koestner, and Richard M. Ryan. 2001. "Extrinsic Rewards and Intrinsic Motivation in Education: Reconsidered Once Again." *Review of Educational Research*, vol. 71, no. 1, pp. 1-27.
- Driver, Julia. 2001. *Uneasy Virtue*. Cambridge: Cambridge University Press.

- 2003. "The Conflation of Moral and Epistemic Virtue." *Metaphilosophy*, vol. 13, no. 3, pp. 367-383.
- Guenin, Louis M. 2005. "Intellectual Honesty." *Synthese*, vol. 145, no. 2, pp. 177-232.
- Hursthouse, Rosalind. 1999. *On Virtue Ethics*. Oxford: Oxford University Press.
- Kamtekar, Rachana. 2004. "Situationism and Virtue Ethics on the Content of Our Character." *Ethics*, vol. 114, no. 3, pp. 458-491.
- LaFollette, Hugh, and George Graham. 1986. "Honesty and Intimacy." *Journal of Social and Personal Relationships*, vol. 3, no. 1, pp. 3-18.
- McMullin, Irene. 2010. "A Modest Proposal: Accounting for the Virtuousness of Modesty." *The Philosophical Quarterly*, vol. 60, no. 241, pp. 787-807.
- Miller, Christian B. 2017. "Honesty." *Moral Psychology, Vol. 5: Virtue and Character*, edited by Walter Sinnott-Armstrong and Christian B. Miller, Cambridge, MA: MIT Press, pp. 237-273.
- Ridge, Michael. 2000. "Modesty as a Virtue." *The American Philosophical Quarterly*, vol. 37, no. 3, pp. 269-383.
- Roberts, Robert C. and W. Jay Wood. 2007. *Intellectual Virtues*. Oxford: Oxford University Press.
- Slote, Michael. 2011. *The Impossibility of Perfection*. Oxford: Oxford University Press.
- Smith, Tara. 2003. "The Metaphysical Case for Honesty." *The Journal of Value Inquiry*, vol. 37, no. 4, pp. 517-531.
- Timpe, Kevin and Craig A. Boyd. 2014. *Virtues and Their Vices*. Oxford: Oxford University Press.

Watson, Gary. 1984. "Virtues in Excess." *Philosophical Studies*, vol. 46, no. 1, pp. 57-74.

Wilson, Alan T. 2016. "Modesty as Kindness." *Ratio*, vol. 29, no. 1, pp. 73-8.

----- Forthcoming. "Avoiding the Conflation of Moral and Intellectual Virtues." *Ethical Theory and Moral Practice*.

Zagzebski, Linda Trinkaus. 1996. *Virtues of the Mind*. Cambridge: Cambridge University Press.